

Chapter 2

Determine Data Needs (Phase II)

Phase II (see Figure 2-1) of the Technical Project Planning (TPP) process is designed to ensure that all data needed to satisfy a site's project objectives are identified. This chapter offers guidance to data users for the detailed level of planning required to determine and document data needed for the current project, and subsequent executable stages. Data users will find guidance in this chapter to help them document their data quality requirements for the intended use(s) of each data need.

Data users must also continue to use their experience, input from others, findings within lessons learned systems, and other technical resources to determine data needs for each site.

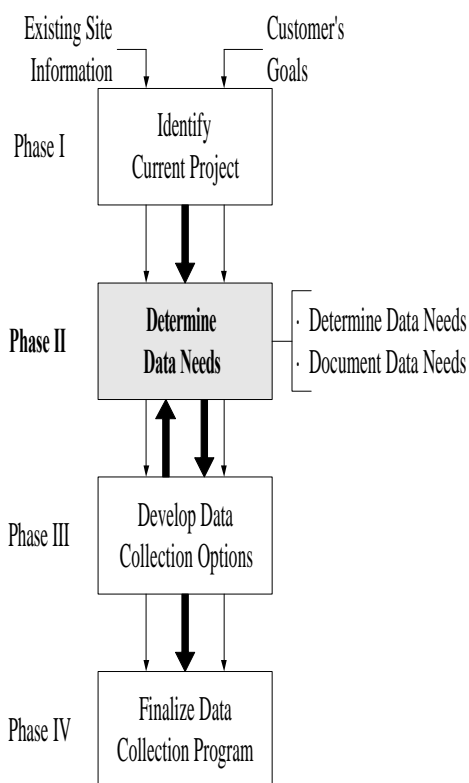


Figure 2-1
Phase II of Four-Phase TPP Process

Data needs determined should include:

- Environmental data needed **from** a site (obtained on-site or by laboratory analysis of a sample from the site); and
- Site information data needed **about** the site (e.g., “as-built” drawings; weather information; water and electrical supply sources; utility conflicts; site access limitations).

2.1 DETERMINE DATA NEEDS.

Determining data needs is an iterative thought process. As presented in this manual, many technical personnel must collaborate to define what is required to satisfy the project objectives.

2.1.1 Review Phase I Memorandum for Record (MFR).

The PM should distribute the Phase I MFR and any project objective worksheets to technical personnel involved in Phase II. Data users' efforts to determine data needs should begin with their review of the Phase I MFR. Review of Phase I information is particularly critical for those personnel not involved in Phase I efforts and for the entire team when some time has passed since Phase I efforts were completed.

2.1.2 Establish Data User's Roles.

Project objectives identified during Phase I should be reviewed to ensure technical personnel understand each project objective. Technical personnel must also be aware of both the “basic” project objectives associated with the current project and those “optimum” project objectives associated with future executable stages. Efforts to establish data user's roles will help focus all technical personnel on their responsibilities and what is required to satisfy the site's project objectives.

EM 200-1-2
31 Aug 98

In most cases, the project manager (PM) should meet with the data users to discuss the preliminary conceptual site model and provide leadership as they discuss what is required to satisfy each project objective. While convened, the TPP team should confirm that they share a common understanding of the preliminary conceptual site model and which data users have a role in determining the data needed to satisfy each project objective.

The PM should also reinforce the premise that data users must work to identify “basic” data needs of the current project; “optimum” data needs that are cost-effective and prudent to fulfill during the current project for a future executable phase; and any “excessive” data needs specifically requested by someone besides the data users, but not needed by the data users.

“Excessive” type data needs should not be created, but identified as a result of data users realizing that some data needs, imposed or mandated by others, are not required to satisfy the basic and optimum project objectives. In some cases, data users will learn that the intended use of the mandated data is actually appropriate, but simply lacked a sufficiently documented project objective. In other cases, data users may realize that the “excessive” data needs imposed by others represent differences in professional opinion or technical judgment as to what data is needed to satisfy a project objective.

Application of the data user perspectives of risk, compliance, remedy, and responsibility will ensure planning is sufficiently detailed to identify the range of data typically required for satisfying project objectives and progressing to site closeout.

2.1.2.1 Risk Data User Perspective.

Risk data users evaluate human health and ecological risks at a site. Technical personnel who collaborate to determine risk-related data needs typically have the following roles at a site:

- Evaluate potential risk-based screening levels to ensure appropriate quantitation limits are established for environmental analyses;
- Perform preliminary determination of hazard or risk to support the decision as to whether further action is warranted;
- Prepare a baseline risk assessment or quantitative evaluation of risk to support a determination of the degree of risk and whether remediation is required;
- Develop remedial action objectives and cleanup levels, as well as detailed analyses of risk reduction provided by remedial alternatives;
- Evaluate suitability of site controls for mitigating short-term risks associated with remediation;
- Verify safety of working conditions for personnel during treatment system construction and operation and maintenance efforts; and
- Evaluate monitoring data to determine the site no longer poses risk and long-term site monitoring can be discontinued.

2.1.2.2 Compliance Data User Perspective.

Compliance data users evaluate and monitor satisfaction of legal and regulatory requirements at a site. Personnel who collaborate to determine legal or regulatory-related data needs typically have the following roles at a site:

- Determine a site’s regulatory compliance with each applicable or relevant and appropriate requirement (ARAR) and later compliance with ARARs;

- Properly manage remediation and investigation derived wastes;
- Contribute to development of remedial action objectives, as well as evaluate remedial alternatives for compliance with each ARAR;
- Verify that implementation of remedial action systems will be compliant with each ARAR;
- Complete procedural requirements under the law governing the response actions (usually CERCLA or RCRA);
- Adhere to the obligations of any agreements, permits, or orders controlling the response actions;
- Determine whether certain comments, requests, or demands from non-federal entities (including regulators), require adherence;
- Predict legal or regulatory issues that will drive response and other regulatory actions; and
- Comply with specific sampling requirements of federal and state programs.

Contaminant fate and transport data needs vary by data user perspective. Therefore, the TPP process encourages each data user perspective to determine their specific fate and transport data needs.

2.1.2.3 Remedy Data User Perspective.

Remedy data users identify possible alternatives for response actions and design all response action components. The role of the remedy data users involves evaluating the feasibility, implementability, or effectiveness of remedies at a site. Remedy data users must also consider potential process interferences and the secondary technologies required to successfully

implement a remedial technology at a site. Technical personnel who collaborate to determine remedy-related data needs typically have the following roles at a site:

- Perform preliminary determination of chemical and physical characteristics of the wastes to begin to determine potential site remedies;
- Identify and screen technologies potentially suitable for mitigating site risks to acceptable levels, as well as perform the detailed analysis necessary to support remedy selection;
- Prepare engineering design and construction plans for response actions, including alternative analysis;
- Optimize operation and maintenance activities and long-term monitoring; and
- Gather cost and performance data needed for life-cycle assessments, evaluation of the technology on similar sites, and incorporation of lessons learned and improvements on future designs.

2.1.2.4 Responsibility Data User Perspective.

Responsibility data users attempt to define what federal or non-federal entity has responsibility for the site's conditions in the event that any response actions are required. Responsibility-related data needs are typically related to determining federal liability at a site, developing a legally defensible position, creating a cost allocation strategy, defining settlement terms with other potentially responsible parties, or presenting or defending in legal proceedings related to responsibility. Some responsibility perspective data needs have elements in common with other data user perspectives (e.g., site history and characterization), and many responsibility perspective determinations are dependent on conclusions of the investigation and design process at a site (e.g., baseline risk

assessment will establish need for response and affect the need for a responsibility determination). Several elements of a responsibility evaluation (e.g., liability determinations, cost allocations) are unique to the responsibility perspective.

2.1.3 Evaluate Use of Existing Data.

Before defining new data needs for a project, data users and data implementors should evaluate the usability of existing data to determine whether additional data are required. Experience has shown that some, if not most, existing data may be suitable for qualitative and for quantitative uses. Detailed usability reviews can determine existing data quality and potential need(s) for additional data to satisfy the project objectives.

The question of whether and how existing data can be used (e.g., in a risk assessment calculation or to support a clean closure) will require specific evaluations of their usability for each intended use. Technical personnel must remember that some existing data may be of an unacceptable quality for one use, but yet of an acceptable quality for another unrelated use at the site.

Review of existing data is a fundamental and critical TPP activity that must occur prior to determining the additional data needed at a site. However, prior to eliminating any data needs from further consideration, the team should be sure that the data user(s) concur that existing data is usable for the intended use(s).

2.1.4 Define Data Needs.

During this TPP activity, technical personnel representing each data user perspective define the data needed to satisfy the project objectives.

Efforts to define data needs must focus on establishing data need requirements for each media type, including sampling areas and depths; chemical concentrations of interest; and the number of samples necessary to satisfy the project objectives.

To identify and organize the data needed, technical personnel should take every advantage of tools such as the preliminary conceptual site model; decision trees or flowcharts; and process diagrams. These tools can provide a logical basis and offer technical personnel a visual prompt for reviewing available site information and defining additional data needs. Other potential tools include data need checklists provided in other technical references. However, data need checklists should not be used as standard lists of data to collect, but as checklists to prompt data users to identify the site-specific data needed to satisfy project objectives at a site. While defining data needs, data users should:

- Consider the consequences of unacceptable decisions or decision errors throughout completion of the work at the site;
- Consider how much data is required;
- Consider data collection approaches, including expedited site characterization and field screening approaches;
- Consider the cost of additional data collection in dollars and time; and then
- Decide how data needs can be balanced within project cost and schedule constraints.

2.1.4.1 Probabilistic/Non-Probabilistic Decisions.

As data users define data needs and the number of samples required, they must recognize that both probabilistic and non-probabilistic data needs should be identified, as appropriate based on intended data uses and the project objectives.

When a data user defines a probabilistic-type data need during this TPP activity, the data user should use Steps 5 and 6 of the U.S. Environmental Protection Agency's (EPA's) 7-Step Data Quality Objective (DQO) Process to determine the number of samples required for the intended data use.¹ (Appendix E presents a detailed "crosswalk" to EPA's 7-Step DQO Process from the TPP process.)

Although powerful, obtaining concurrence among decision makers regarding probabilistic decisions can be difficult. Application of probabilistic methods can only be accomplished when these three conditions exist:

- (1) A precise study question is defined;
- (2) The customer and lead regulator are willing to and successful in establishing tolerable limits on decision errors; and
- (3) The support of a qualified environmental statistician is available to work on the project.

When probabilistic methods are either inappropriate or cannot be employed for a data need because the three conditions do not exist, data collection planning can be judgmentally based on the expertise of the technical personnel representing the applicable data user perspective.

2.1.4.2 Number of Samples.

Each data user is responsible for identifying the number of samples, or decision logic, required for each data need based on the intended data use(s) and the project objectives.

When non-probabilistic or judgmental sampling is appropriate, the number of samples may be designated by guidance or technical literature specific to the data user perspective. In some cases, the number of samples needed to satisfy an objective (e.g., determining if a contaminant is present) may be based on experienced judgment of the technical personnel representing the data user perspectives for each specific site.

In some instances, data needs should be fulfilled using probabilistic or random sampling where the number of samples required to make the related site decision should be calculated. For example, project objectives that have specific data needs (e.g., determining if the contaminant levels detected are sufficiently different from the background levels of the constituent at the site) may provide a suitable opportunity to use a statistical basis to establish the number of samples to be collected. In any case, it is important that data users recognize that use of statistical techniques as the basis for designing environmental sampling plans can reduce the number of unnecessary samples taken in the field, and improve the sampling representativeness by quantifying the statistical uncertainty of the sampling design. Inappropriate application of statistics for probabilistic data needs can also result in either the collection of too many or too few samples.

When necessary, in accordance with recommendations within EPA's Guidance for Data Usability for Risk Assessment, risk assessment personnel should indicate the number of samples in terms of classical

statistics.¹³ The remedy perspective, on the other hand, typically uses engineering judgment or other performance criteria as a means to designate the number of samples required to support a remedy-related data need.

Decisions to use classical statistics methods or geostatistical methods must also be based on the intended data use(s) and known or anticipated variability of the data in the environment. This is the case because randomly distributed variables or data are suited for classical statistics applications and spatially related or regionalized variables that have continuity from point to point present opportunities to use geostatistics.

2.1.4.2.1 Applications of Classical Statistics. Classical statistics tools (e.g., random, stratified random, or systematic random sampling designs) can be used to determine the number of samples required to support various probabilistic decisions. Classical statistics can be used to determine the number of samples required to define representative concentration values (e.g., background soil concentrations) or evaluate trends (e.g., waste pile sampling, chemical concentrations in soils) over an area of interest. Classical statistics are most appropriate for mean concentrations; however, other methods may be more appropriate or suitable for comparing populations or identifying “hot spots.” Classical statistics methods can be used to determine the number of samples needed from each medium (or each stratum within a medium) to provide sufficient data to support project objectives.

2.1.4.2.2 Applications of Geostatistics. Geostatistics are a specific branch of statistics used to optimize the co-variance of a variable of interest and can involve classical (or simple) random, stratified random, or systematic random sampling designs. Geostatistical

techniques are appropriate for environmental sampling programs intended to define or evaluate the distribution of contamination at a site or within an area of interest. Geostatistics are particularly useful for identifying “hot spots” and calculating the reasonable maximum exposure for risk assessments. It can also be used to produce probability estimates of a variable of interest based on recognized geostatistics methods (e.g., semivariogram analysis, cross-validation, and data kriging). Data users and data implementors should seriously consider the use of geostatistical methods since they can provide considerable support to the development of data collection programs and result in significant cost savings. Potential applications of geostatistics include:

- Contour mapping and interpolation;
- Identification of sample locations;
- Optimization of sampling existing monitoring devices;
- Risk assessment/probability estimation; and
- Remedial design.

2.1.4.3 Data Collection Considerations.

While defining data needs, each data user should re-evaluate earlier considerations of using either field screening or expedited site characterization (ESC) approaches for collecting site data. As discussed in Paragraph 1.3.2.4, the team may chose to conduct some field screening activities concurrent with Phase I, II, or III TPP activities to refine their understanding of a site. Use of an ESC approach typically expects data users to first establish a site’s physical setting before contaminant investigations are conducted as discussed in Paragraph 1.3.2.5.

When data users have identified appropriate opportunities to use field screening or ESC approaches, they should advise the PM and data implementors which data needs are candidates for using either approach. In those instances

where field screening or ESC approaches will be used, it will even be more critical for the data users to provide the decision logic information that can be incorporated into the corresponding dynamic work plan for the site. Data users will also need to provide a description or decision flowchart of the rationale that should be used for making field decisions contingent on the results of previous samples.

Data users must also recognize that data needs identified during this TPP activity should include both site information and environmental data. Appendix F provides a site information worksheet and several data need worksheets that are recommended for documenting the data needs of the data users.

When defining each data need, data users are responsible for communicating whether a data need is a “basic” data need that contributes to satisfying a current project objective, an “optimum” data need that would be cost-effective and prudent to fulfill during the current project, or an “excessive” data need specifically requested by someone other than the data users, and not needed by the data users. (Paragraph 3.2 further describes the data collection options of “basic,” “optimum,” and “excessive”.)

2.1.4.4 Risk Data Needs.

Using the preliminary conceptual site model developed during Phase I, the risk data users should conceptualize and identify the data needed to address each of the pathways that will be part of the risk assessment for the site. In assessing risks to human and environmental receptors, one must be able to show a relationship between potentially exposed populations (for both current and future site use) and the chemicals detected onsite. This relationship is evident in the elements of a risk assessment (e.g., conceptual site models; data

evaluation and chemical of potential concern identification; exposure assessment; toxicity assessment; and risk characterization).

Future land use pathways (if different from current pathways) will require data to support transport models suited for evaluating spatial and temporal behavior of the chemical(s) at the site over time. Therefore, the risk perspective must determine the most appropriate models to satisfy the project objectives since data requirements vary by model.

2.1.4.5 Compliance Data Needs.

Compliance data users should compare site conditions or activities with legal and regulatory requirements and standards to establish the governing laws and regulations and to determine what is required for site compliance. They must also compare possible site conditions or activities that are regulated (e.g., treatment, storage, and disposal) with applicable regulatory standards. Potentially applicable regulatory standards are defined by the primary regulatory program and may specify chemical analysis requirements and point(s) of compliance (location and type of samples) used to assess compliance. Compliance data user efforts to define compliance data needs should involve:

- Review of the project objectives identified from the primary governing statutes (e.g., CERCLA Sections 104, 120, 121; RCRA Sections 3004u and 3008h) and the applicable regulations;
- Identification of activities or conditions that give rise to certain standards, requirements, or criteria that must be satisfied (e.g., treatment, storage, and disposal; drinking water contamination; surface water discharge);
- Consideration of potentially affected media (i.e., air, surface water, sediment, soil, groundwater);

- Identification of chemical-, action-, and location-specific ARARs;
- Identification of point(s) of compliance (e.g., drinking water aquifer, effluent discharge, stack emissions);
- Compilation of documents, reports, data, correspondence, etc., that demonstrate satisfaction of procedural requirements arising from laws, regulations, agreements, permits, or orders; and
- Identification of community relations and public involvement activities for outreach to interested stakeholders.

Compliance data needs will be both qualitative and subjective (point of compliance), as well as quantitative (environmental data needs).

2.1.4.6 Remedy Data Needs.

Remedy data users define data needed to identify, screen, and analyze possible response action alternatives at a site. The efforts to define remedy data needs will depend on the phase of a site's progress to site closeout. Remedy data needs become more complex as the alternative evaluation process proceeds from technology identification to remedy selection and design, and finally operation and maintenance of treatment systems.

During the early stages of a site's progress to site closeout, technical personnel should begin to consider possible general technologies that may be applicable to the site (e.g., containment; excavation and disposal; in-situ treatment). Site information type data needs are typically sufficient to support these evaluations (e.g., contaminant characteristics, physical characteristics of the site, and physical features of the site). The next level of evaluation includes identification of common technologies (e.g., soil washing, incineration, capping) that relate to the general technology type(s)

previously identified for a site. The remedy data needs for technology screening are typically environmental type data needs (e.g., soil moisture content, pneumatic permeability, and cation exchange capacity). Based on technology screening results, only a few alternatives for remedy selection and design are further considered. While constructing, operating, maintaining, and monitoring a remedy, ongoing efforts will be expended by the remedy data users to obtain cost and performance information for optimizing the treatment system and for similar systems in the future. The remedy data needs to support these later evaluations will be more complex and require both site information and environmental data (e.g., treatability studies, soil compaction, and available water sources).¹⁴

Remedy design-type personnel should contribute to TPP efforts beginning with the earliest stages of site assessment and investigation. Remedy construction-type personnel should begin to contribute to TPP efforts when site remedy selection and design activities begin.

2.1.4.7 Responsibility Data Needs.

The technical and legal counsel personnel responsible for defining responsibility data needs will not only be concerned with determining the legal basis for a response action, but also with defining responsibility at a site. Responsibility data users must rely on legal counsel to identify the phase of execution and specific position and negotiation strategies that will affect the identification of responsibility data needs.

For example, one emphasis would be to obtain data for determining a site's eligibility under the Formerly Used Defense Site program and identifying the potential for another potentially

responsible party (PRP).¹⁰ In this case, responsibility data would need to be collected toward the goal of settling with the other PRP. In another instance, responsibility data needs would involve collecting past disposal records for position development purposes that ultimately contribute to developing a cost allocation formula during negotiations with other PRPs.

Background and historical site information will make up much of the responsibility data needed to develop a negotiation position. This includes articles of incorporation; facility ownership records; contract documents; lease agreements; historic process and operations information; federal and industry information on standard practices related to the chemicals of concern; manifests; disposal logs; and aerial photos. The site characterization data can focus the historical research toward the use and disposal of specific chemicals at specific locations based on observed contamination. Historical information should guide site characterization work by narrowing the list of analyses and general sampling locations. These two investigations should be done in parallel to be successful.

2.2 DOCUMENT DATA NEEDS.

Personnel representing data user perspectives are responsible for communicating their data needs so the needs can be incorporated within data collection options developed during Phase III activities. Communicating or documenting data needs are critical TPP activities that lead to successful project execution. Documenting data needs, as discussed here, is the recommended means for technical personnel to communicate their data needs.

This manual offers several options for documenting data needs given the wide range of data needs and data uses. Appendix F offers a

site information worksheet and a series of data need worksheets for documenting data needs of the risk, compliance, remedy, and responsibility data user perspectives. Use of standardized data need worksheets will allow quick and easy quality assurance/quality control review of the data need planning.

The critical aspects of documenting data needs can be reduced to the following.

- What data is needed (e.g., contaminant or characteristic of interest, and media)?
- Who needs the data (i.e., risk, compliance, remedy, or responsibility data user perspective)?
- What is the intended data use(s) (e.g., contaminant fate and transport; baseline risk assessment; remedial design; operation and maintenance plan) to satisfy project objectives?
- What number of samples are required to satisfy the intended use(s), including whether the number of samples is fixed, somehow contingent upon field screening results, or is the minimum anticipated by the dynamic decision logic approach defined by the data user?
- What is reference concentration of interest or other performance criteria (e.g., action level, compliance standard, decision level, design tolerance)?
- Where is area of interest or desired sampling location(s) and depth(s)?

Site information worksheets and data need worksheets are comprehensive lists of the data needs at a site. Information presented on the worksheets identify additional data needed by each data user perspective to satisfy the project objectives.

2.3 COMPLETE PHASE II ACTIVITIES.

The technical personnel should review the data need worksheets to ensure that each project objective has been considered and related data need considerations have been made by each applicable data user perspective. In accordance with the applicable quality management plan, the PM should also have independent technical resources review the data need worksheets. (The data need worksheet examples provided in Appendix F may be useful during independent review efforts.) In any case, all projects will be periodically evaluated by the project team to ensure baseline requirements of scope, schedule, and cost are being met.²

If it appears that some project objectives have no associated data needs, the PM should meet with the technical personnel and confirm that no additional data is needed to support the particular project objectives. The PM or technical personnel should document in the project file why specific project objectives require no additional data. The PM should also meet with the technical personnel to understand any instances when no “optimum” data needs have been identified during Phase II activities.

The PM should review any site information worksheets or lists of site information data needs that have been identified by the data users. It is the PM’s responsibility, working with the technical personnel, to decide how and when site information needs will be fulfilled (e.g., discussions with the customer, site visits, incorporated within appropriate scope of work or work plan sections).

The PM should then distribute copies of all data need worksheets and any attached illustrations to all appropriate TPP team members.